Strong localization of positive charge in DNA induced by its interaction with environment

A. L. Burin and D. B. Uskov

Department of Chemistry, Tulane University, New Orleans LA, 70118

(Dated: June 6, 2008)

Abstract

We investigate a quantum state of positive charge in DNA. A quantum state of electron hole is determined by the competition of the pi-stacking interaction b sharing a charge between different base pairs and the interaction λ with the local environment which attempts to trap charge. To determine which interaction dominates we investigate charge quantum states in various $(GC)_n$ sequences choosing DNA parameters satisfying experimental data for the balance of charge transfer rates $G^+ \leftrightarrow G_n^+$, $n = 2, 3$. We show that experimental data can be consistent with theory only assuming $b \ll \lambda$ meaning that charge is typically localized within the single G site. Consequently as follows from our modeling consideration any DNA duplex including the one consisting of identical base pairs cannot be considered as a molecular conductor. Our theory can be verified experimentally measuring balance of charge transfer reactions $G^+ \leftrightarrow G_n^+$, $n \geq 4$ and comparing the experimental results with our predictions.

PACS numbers: 7080.Le, 72.20.Ee, 72.25.-b, 87.14.Gg

I.   INTRODUCTION

Positive charge (hole) transfer in DNA is extensively investigated since its experimental discovery [1]. This process can be responsible for the oxidative DNA damage [1, 2, 3, 4, 5] and is possibly important for DNA repairing [6, 7]. Also an ability of DNA to promote long distant charge transfer can be used in molecular electronics applications [8]. These studies raise the question, whether DNA is a molecular conductor. In molecular conductors (e. g. carbon nanotubes) charge is usually delocalized within several sites (monomers). Such system possesses a high conductivity similarly to metals. The alternative behavior takes place when the charge is typically localized within the single site and it hops to adjacent sites due to rare environment fluctuations. Based on the results of direct measurements of charge transfer in DNA [9] we argue in this paper that DNA is not a molecular conductor.

DNA contains two different sorts of base pairs AT and GC forming quasi-random sequences. The lowest ionization potential is attributed to a GC pair (essentially G-base [10]). Since the electron transfer integral b between adjacent bases does not exceed AT - GC ionization potential difference $\Delta \sim 0.5$eV, the quantum state of charge in a static environment will be localized near some G base and its localization length is comparable to the interbase distance [11, 12]. Interaction with environment breaks down this localization, inducing charge hopping between quantum states localized at adjacent G sites. This happens because of rare environment fluctuations supporting delocalization of charge between neighboring G bases. Indeed, according to experimental studies [4, 5, 9] and theoretical models [13, 14] the sequence dependent charge transfer in DNA can be represented as the series of charge hops between adjacent G bases serving as centers of localized states. An addition of AT pair between GC base pairs dramatically reduces the charge hopping rate [4, 13, 14] and therefore the optimum base sequence for the efficient charge transfer consists of identical base pairs. We are going to study a charge (hole) quantum states in sequences of identical GC pairs of various lengths, as most promising candidates for efficient charge transfer.

The thermal energy at room temperature $k_B T \sim 0.026$eV is very small compared to at least one of the characteristic energies of the system including the reorganization energy λ and the electron transfer integral b. The reorganization energy is caused by the electrostatic interaction of the hole with the environment and its minimum estimate is $\lambda \sim 0.25$eV

[15, 16, 17, 18, 19]. The electron transfer integral b is associated with the pi-stacking interaction of heterocyclic groups in DNA bases. It is defined as the gain in the hole energy due to its sharing between two adjacent DNA bases having equal energies with respect to the energy of the hole localized within a single G base [18]. The minimum estimate for the electron transfer integral between adjacent G bases is around 0.1eV [18]. Thus the thermal energy is, indeed, small compared to other characteristic energies and we can use the ground state of the hole coupled to environment as the representative state. One should notice that the ground state can be used as a representative state, but it strongly differs from transition states responsible for the charge transfer. Since the hole transitions are quite rare the "typical" charge state can be described ignoring them.

The spatial size of the hole ground state is determined by the competition of the delocalization of charge due to the pi-stacking interaction of heterocyclic groups belonging to adjacent bases and the localization caused by the environment polarization around the charge. The charge delocalization energy is characterized by the effective electron transfer integral b and the localization energy is given by the environment reorganization energy λ [15, 16, 17, 18, 19]. Delocalization of charge over k base pairs leads to the gain in the kinetics energy $E_{del} \sim -2b + b/k^2$, similarly to the classical problem of particle in the box. The reorganization energy scales with the size of charge wavefunction as $E_{loc} \sim -\lambda/k$ (see e. g. [15, 20]). This dependence can be understood as following. The local environment polarization induced by the hole is proportional to the charge density $P \propto d \propto 1/k$, where d stands for the local charge density. Consequently the interaction in each site having charge e/k with the environment polarization is proportional to their product $pd \propto 1/k^2$ and the sum of all these k interactions over all k sites results in the dependence λ/k. This dependence takes place only in a one-dimensional system. Since at large k the reorganization energy $-\lambda/k$ dominates over the kinetic energy b/k^2, the ground states is always localized and the localization radius can be estimated using the number of sites k minimizing the total energy

$$E_{tot}(k) \approx -2b + \frac{b}{k^2} - \frac{\lambda}{k}. \tag{1}$$

The minimum takes place at $k = k_* \approx 2b/\lambda$ which is the number of DNA bases occupied by the hole in its ground state. At zero temperature the hole is localized, while at finite temperature it can hop between different states because of its interaction with the fluctuating environment. In the translationally invariant system ($(GC)_n$ or $(AT)_n$) the potential barrier

separating two configurations can be estimated as the energy price for the environment fluctuation increasing the size of the wavefunction by one more site compared to its optimum state $k_* \rightarrow k_* + 1$, which can be expressed as $E_{tot}(k_* + 1) - E_{tot}(k_*)$ Eq. (1). Such fluctuation creates the transition state and after the relaxation of environment the hole can arrive at the new equilibrium centered at different G base. If $b \ll \lambda$ then this activation energy is given by approximately a quarter of the reorganization energy $\lambda/4 \gg k_B T$. Then we expect the hole mobility to obey the Arrhenius law as in insulators. In the opposite limit we obtain a very small value $\Delta \sim \lambda^4/(16b^3)$, which becomes negligible at moderately large b/λ. For instance if $b = 1$eV the potential barrier for the hole transition becomes smaller than the thermal energy $k_B T \approx 0.026$eV already at $\lambda < 0.7eV$. In this regime charge transport is weakly sensitive to the temperature as in metals.

Therefore it is very important to determine the true relationship of b and λ in DNA. This relationship characterizes its conducting behavior. Indeed, in the regime $b > \lambda$ a DNA molecule made of identical base pairs would behave as a one-dimensional conductor, while in the opposite limit DNA cannot be treated as a molecular wire. Existing estimates in literature [10, 15, 16, 17, 18, 19, 20] do not help much in determining the relationship of b and λ, because there is a large controversy between different studies. In particular, various estimates for the electron transfer integral b range from 0.05eV [18] to 0.5eV [10]. Also all calculations of b ignore vibrational rearrangements. Polar vibrational modes associated with the covalent bonds possess a high vibrational quantum energy $\hbar\omega \sim 0.3$eV $\gg k_B T$. At room temperature these modes must remain in the ground state during charge tunneling. However since the tunneling of hole changes the equilibrium coordinates of polar vibrations they must tunnel together with the hole. This leads to the redefinition of the effective tunneling amplitude b (if it is smaller then the reorganization energy λ_v) associated with the given vibrational mode [22, 23, 24] as

$$b_{eff} = b \cdot e^{-\frac{\lambda_v}{2\hbar\omega}}. \tag{2}$$

To our knowledge this effect is ignored in all existing theoretical estimates of the electron transfer integral.

The calculations for the reorganization energy λ associated with the classical interaction with solvent were made using the continuous medium approach. The estimates of λ range from 0.25eV [19] to more than 1eV [15] due to uncertainty in the water dielectric constant

value near the DNA molecule. It is hard therefore to select the right parameters based on existing theoretical results.

We suggest an alternative method to study the hole quantum state within the DNA molecule, using experimental data sensitive to the relationship of two key parameters of the theory b and λ. Namely, we exploit the rate constants for the balance of charge transfer reactions between different $(GC)_n$ complexes, measured by Lewis and coworkers [9]

$$
\begin{aligned}
G^+ + GG &\underset{k_{-t}^{GG}}{\overset{k_t^{GG}}{\rightleftharpoons}} G + (GG)^+, \quad \frac{k_t^{GG}}{k_{-t}^{GG}} = 7.7 \pm 1 \\
G^+ + GGG &\underset{k_t^{GGG}}{\overset{k_t^{GGG}}{\rightleftharpoons}} G + (GGG)^+, \quad \frac{k_t^{GGG}}{k_{-t}^{GGG}} = 20 \pm 1.
\end{aligned}
\tag{3}
$$

In the thermal equilibrium these ratios are determined by base pair partition functions

$$
\begin{aligned}
r_2 &= \frac{k_t^{GG}}{k_{-t}^{GG}} = \frac{Z_{2+}Z_1}{Z_2 Z_{1+}}, \\
r_3 &= \frac{k_t^{GGG}}{k_{-t}^{GGG}} = \frac{Z_{3+}Z_1}{Z_3 Z_{1+}}.
\end{aligned}
\tag{4}
$$

where Z_{n+} stands for the partition function of G_n sequence containing the single hole in it, while Z_n is the partition function of the same base sequence, but without the hole. Since G and G_n complexes are separated by some AT bridge we can ignore their interactions.

We assumed that the charge transfer between G and G_n complexes is very slow compared to the charge transfer rate inside the complex. This agrees with experiments showing the strong reduction of the charge transfer rate with the length of AT separating GC pairs [4]. Consequently the contribution of highly excited states like the transition states for the charge transfer between complexes to the partition functions can be neglected so we can assume that the local equilibrium of charge within G_n sequence is established. Accordingly one can use Eq. (4) for the ratio of charge transfer rates.

Both ratios in Eq. (3) are determined by parameters b and λ and the thermal energy at room temperature $k_B T \sim 0.026 eV$. Below we calculate both ratios using tight binding model for G_n complexes and standard linear response theory for charge interaction with the environment [16, 21]. Theory determines the domain of parameters λ and b satisfying experimental data Eq. (3). *We demonstrate that any reasonable choice of λ and b, satisfying Eq. (3), corresponds to the regime $b \ll \lambda$, where the hole in its ground state is localized essentially in a single G base (see Fig. 1).*

The manuscript is organized as following. In Sec. II we formulate and discuss the model for the hole in G_n sequence coupled to the classical environment. In Sec. III we derive expressions for rate ratios r_n. In Sec. IV A we discuss the properties of the hole ground state in the domain of its parameters consistent with the experiment [9] and show that it is localized essentially within the single G site. In Sec. IV B we discuss the partition functions of charge using the simple perturbation theory approach and predict the balance of charge transfer reactions between G^+ state and G_n^+ state for arbitrarily number n of G bases. In Sec. IV C we discuss the effect of correlations of environment polarizations on different G sites on ratios r_n. In Sec. IV D we investigate the influence of A bases surrounding G_n sequences in the experiment [9] on our conclusions. In Sec. V we resume our results and discuss the experiments, which can verify our theory and help to determine parameters b and λ more accurately.

II. MODEL

The chain of n GC base pairs can be described by the tight binding Hamiltonian coupled to the classical environment represented by coordinates X_i, $i = 1, ...n$ for each DNA G site

$$\widehat{H} = \widehat{H}_{hole} + \widehat{H}_{int},$$

$$\widehat{H}_{hole} = -b \sum_{i=1}^{n-1} (c_i^+ c_{i+1} + c_{i+1}^+ c_i),$$

$$\widehat{H}_{int} = \frac{1}{2\lambda} \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i n_i,$$

$$n_i = c_i^+ c_i. \tag{5}$$

Here c_i, c_i^+ are operators of creation and annihilation of electron hole in a site i. Classical coordinates X_i describing the polar environment, including solvent and counterions, are directly coupled to the local charge density $n_i = c_i^+ c_i$ and the parameter λ stands for the reorganization energy associated with the interaction of charge and its environment. The environment energy is expressed as a bilinear form with respect to solvent coordinates, which is justified by a standard assumption that polarization fields are small compared to atomic fields [21] so we ignore X^3 terms. Since the potentials X include the contribution of counterions it is not clear whether the ion effect can be treated in the linear response approximation. However, we do not think that the contribution of ions can critically change our

assumptions because as was demonstrated by Voityuk and coworkers [25] the contribution of ions does not exceed 20% of the overall charge energy fluctuations, while the remaining 80% of fluctuation are associated with water. As discussed below in Sec. IV B the validity of the linear response approximation does not crucially affect our consideration in the strong localization regime.

We assume that only classical degrees of freedom with excitation energy comparable or less than the thermal energy are left in Eq. (5), while high energy modes are integrated out. This may lead to the renormalization of parameters in the system Hamiltonian Eq. (5) (see e. g. Eq. (2) and Refs. [22, 23, 24]) and we assume that this renormalization is made. We do not include off-diagonal terms $X_i X_j$, $i \neq j$ into the Hamiltonian. This is somehow justified because they are smaller than the diagonal ones [17]. We will show below that for G_2 sequence the problem including off-diagonal terms can be reduced to the diagonal model Eq. (5) with the replacement of the single site reorganization energy λ with the reorganization energy for charge transfer between adjacent sites. For GGG sequence the similar replacement with removal off-diagonal terms remains a good approximation (see Sec. IV C). Therefore all our results are valid if we consider the more general interaction model with the replacement of the single site reorganization energy λ with the charge transfer reorganization energy λ_* as described below in Sec. III.

We assume the electron transfer integral b to be independent of the environment fluctuations. Fluctuations of the electron transfer integral were treated as less significant compared to fluctuations in local site energies because the change of the site energy by more than the thermal energy strongly modifies the tunneling rate, while the change in the electron transfer integral requires the energy fluctuation comparable to the energy $\delta E \sim \hbar/\tau$ where τ is the tunneling time for the electron transition [26]. This energy δE can be comparable to the barrier height [27] which is much larger than the thermal energy. The above expectation conflicts with the molecular dynamics simulations of the electron transfer integral between hole states in adjacent G bases affected by the interaction with environment (see Ref. [28] and references therein), where the remarkable effect of fluctuations on the electron transfer integral has been found. However, molecular dynamics results should be considered with care because of the classical treatment of the nuclei motion, which can be partially of the quantum mechanical nature. We estimated the effect of the distribution of the electron transfer integral b on the ratios of charge transfer rates for the system described by Eq.

(5). The distribution of b was approximated by the Gaussian distribution with parameters $< b >\approx 0.046eV$ and $\delta b \approx 0.064eV$ [28] while reorganization energy was taken as $\lambda = 0.24eV$ [25]. The calculations result in the overestimated ratio of charge transfer rates r_2 Eq. (4) by the factor of five. This result queries the accuracy of molecular dynamics approximation to fluctuations. Also the measurements of charge transfer efficiency between G and GGG sequences separated by n AT pairs [29] can be successfully interpreted assuming that the charge transfer integral is constant at least for $n < 5$ AT pairs [30].

In the experiment [9] G_n sequence was surrounded by A bases. The addition of A bases surrounding G_n sequences to our model Eq. (5) can change our results. The changes are associated with the possible sensitivity of the hole energy in the G site to its neighbors and the ability of the hole to come virtually to the adjacent A site. According to quantum chemistry calculations the ionization potential of the G base depends on surrounding bases [31]. However, the difference of ionization potentials for GG^+G, AG^+G and GG^+A does not exceed 0.1eV. It is remarkably smaller than the reorganization energy and therefore we will ignore it. The error of semiempirical methods used to calculate the ionization potential is comparable to the maximum calculated effect $\sim 0.1eV$ so we cannot consider the energy change 0.1eV or smaller. Screening of electrostatic interaction by water can also reduce the effect of neighbors on the G^+ state ionization potential.

The changes associated with the extension of the charge wavefunction to adjacent A sites can be more important. However these changes do not affect our main conclusion about the strong localization of the charge ground state as shown in Sec. IV D. We ignore the second strand (C_n) because of the weak coupling between strands [18] and the large difference of G and C ionization potentials exceeding 1eV [10].

III. EVALUATION OF RATE RATIOS

We study the ratios of charge transfer rates Eq. (4). Each partition function is given by $Z_n = \int dX_1...dX_n Tr e^{-\beta \widehat{H}_n}$, where $\widehat{H}_n$ is the Hamiltonian Eq. (5) describing G_n sequence, trace is taken only over states with the single hole (Z_{n+}) or no holes (Z_n) and $\beta = 1/(k_B T)$. If there is no hole calculations are reduced to multiple evaluation of a Gaussian integral

8

leading to the expression

$$Z_n = c^n, \quad c = \sqrt{\frac{2\pi\lambda}{\beta}}. \tag{6}$$

For the sequences containing a hole an analytical expression can be obtained only for $n = 1$

$$Z_{1+} = ce^{\beta\lambda/2}. \tag{7}$$

For any number n of GC pairs calculations can be simplified integrating the partition function over a "center of mass" coordinate $X_1 + ..X_n$, which is coupled to the conserving operator of the total number of particles $c_1^+ c_1 + .. + c_n^+ c_n = 1$. In the case $n = 2$ one can conveniently introduce the new coordinates as $X = X_1 + X_2$ and $u = X_1 - X_2$. Then the Hamiltonian of two base pairs (Eq. (5) for $n = 2$) takes the form

$$\widehat{H}_2 = \frac{X^2}{4\lambda} - \frac{X}{2} + \frac{u^2}{4\lambda} - \frac{u(n_1 - n_2)}{2} - b(c_1^+ c_2 + c_2^+ c_1). \tag{8}$$

Eigenenergies of this Hamiltonian can be written as

$$E_{\mp} = \frac{X^2}{4\lambda} - \frac{X}{2} + \frac{u^2}{4\lambda} \mp \sqrt{\frac{u^2}{4} + b^2}. \tag{9}$$

Accordingly the partition function takes the form

$$Z_{2+} = \frac{1}{2} \int_{-\infty}^{+\infty} dX e^{\frac{-\beta X^2}{4\lambda} + \frac{\beta X}{2}} \int_{-\infty}^{+\infty} du e^{-\frac{\beta u^2}{4\lambda}} 2\cosh\left(\beta\sqrt{\frac{u^2}{4} + b^2}\right), \tag{10}$$

The factor $1/2$ is concerned with the coordinate transformation. Performing integration over the center of mass coordinate X we get

$$Z_{2+} = (\sqrt{2}c)e^{\beta\lambda/4} \int_{-\infty}^{+\infty} du e^{-\frac{\beta u^2}{4\lambda}} \cosh\left(\beta\sqrt{\frac{u^2}{4} + b^2}\right). \tag{11}$$

The ratio r_2 of reaction rates Eq. (4) takes the form

$$r_2 = \sqrt{\frac{\beta}{\pi\lambda}} e^{-\beta\lambda/4} \int_{-\infty}^{+\infty} du e^{-\frac{\beta u^2}{4\lambda}} \cosh\left(\beta\sqrt{\frac{u^2}{4} + b^2}\right). \tag{12}$$

Consider GG sequence with off-diagonal interaction of environment coordinates

$$\widehat{H}_{env} = \frac{A_0(X_1^2 + X_2^2)}{2} + A_1 X_1 X_2. \tag{13}$$

9

The coefficients $\widehat{A}$ can be expressed in terms of the average coordinate fluctuation matrix $\widehat{B}$ as

$$\begin{pmatrix} A_0 & A_1 \\ A_1 & A_0 \end{pmatrix} = \begin{pmatrix} B_0 & B_1 \\ B_1 & B_0 \end{pmatrix}^{-1}, \tag{14}$$

where the elements of the matrix $\widehat{B}$ are defined as [17]

$$k_B T B_0 = <X_1^2> = <X_2^2>, \ \ k_B T B_1 = <X_1 X_2>. \tag{15}$$

$< ... >$ means the standard thermodynamics averaging with the Hamiltonian Eq. (13).

Then we can still proceed to the new coordinates X and u defined as previously. The Hamiltonian can be written as

$$\widehat{H}_{p1} = \frac{(A_0 + A_1)X^2}{4} - \frac{X}{2} + \frac{(A_0 - A_1)u^2}{4} - \frac{u(n_1 - n_2)}{2} - b(c_1^+ c_2 + c_2^+ c_1). \tag{16}$$

Note that the Hamiltonian for the single site has different definition of the diagonal term of the environment energy ($n = 0, 1$ is the population of the single G site)

$$\widehat{H}_s = \frac{X^2}{2B_0} - nX. \tag{17}$$

This is because the Hamiltonian must provide the right expressions for average squared polarizations in Eq. (15). The partition functions involved in our consideration for r_2 can be expressed as

$$Z_1 = \sqrt{\frac{2\pi B_0}{\beta}}, \ \ Z_{1+} = \sqrt{\frac{2\pi B_0}{\beta}} e^{\beta B_0/2}, \ \ Z_2 = \frac{2\pi}{\beta \sqrt{A_0^2 - A_1^2}},$$

$$Z_{2+} = \sqrt{\frac{\pi(B_0 + B_1)}{\beta}} e^{\beta(B_0 + B_1)/4} \int_{-\infty}^{+\infty} du e^{-\frac{\beta u^2}{4\lambda_*}} 2\cosh\left(\beta\sqrt{\frac{u^2}{4} + b^2}\right),$$

$$\lambda_* = B_0 - B_1. \tag{18}$$

The expression for the ratio r_2 Eq. (4) takes the form

$$r_2 = \sqrt{2\frac{\beta}{\pi(B_0 - B_1)}} e^{-\beta(B_0 - B_1))/4} \int_{-\infty}^{+\infty} du e^{-\frac{\beta u^2}{4(B_0 - B_1)}} \cosh\left(\beta\sqrt{\frac{u^2}{4} + b^2}\right). \tag{19}$$

This expression has the form identical to Eq. (12) if the reorganization energy λ is replaced with $\lambda_* = B_0 - B_1$. The new parameter λ_* is the reorganization energy for the charge transfer between adjacent sites. This is the relevant parameter which can be used also for GGG sequence and more complicated sequences as it will be shown in Sec. IV C.

The expression for r_3 is more complicated. Below we give the result after integration over the coordinate of the center of mass $X = X_1 + X_2 + X_3$ and use two other coordinates $v = X_1 - 2X_2 + X_3$ and $u = X_1 - X_3$ in the case of the "diagonal" environment energy Eq. (5)

$$r_3 = \frac{\beta}{4\sqrt{3}\pi\lambda} e^{-\frac{\beta\lambda}{3}} \int_{-\infty}^{+\infty} dv \int_{-\infty}^{+\infty} du\, e^{-\frac{\beta v^2}{12\lambda} - \frac{\beta u^2}{4\lambda}} \cdot Tr\left(\exp(-\beta\widehat{V}_3)\right),$$

$$\widehat{V}_3 = \begin{pmatrix} \frac{v}{2} + \frac{u}{6} & -b & 0 \\ -b & -\frac{u}{3} & -b \\ 0 & -b & -\frac{v}{2} + \frac{u}{6} \end{pmatrix}. \tag{20}$$

Similarly one can obtain the expression for r_4

$$r_4 = \frac{1}{4}\left(\frac{\beta}{2\pi\lambda}\right)^{\frac{3}{2}} e^{-\frac{3\beta\lambda}{8}} \int_{-\infty}^{+\infty} dt \int_{-\infty}^{+\infty} dv \int_{-\infty}^{+\infty} du\, e^{-\frac{\beta t^2}{8\lambda} - \frac{\beta u^2}{4\lambda} - \frac{\beta v^2}{4\lambda}} \cdot Tr\left(\exp(-\beta\widehat{V}_4)\right);$$

$$\widehat{V}_4 = \begin{pmatrix} \frac{t}{4} + \frac{u}{2} & -b & 0 & 0 \\ -b & -\frac{t}{4} + \frac{v}{2} & -b & 0 \\ 0 & -b & -\frac{t}{4} - \frac{v}{2} & -b \\ 0 & 0 & -b & \frac{t}{4} - \frac{u}{2} \end{pmatrix};$$

$$t = X_1 - X_2 - X_3 + X_4,\ u = X_1 - X_4,\ v = X_2 - X_3. \tag{21}$$

Eqs. (12), (20), (21) were used for the numerical evaluation of partition functions and ratios r_2, r_3 and r_4.

IV. DEFINITION OF PARAMETERS CONSISTENT WITH EXPERIMENT

A. Numerical results

We have performed numerical evaluations of ratios in Eq. (4) to find domains of parameters b and λ satisfying Eq. (3) and show these domains in Fig. 1. The upper (lower) border of each domain is defined by the maximum (minimum) value of ratios r_2 and r_3 Eq. (4) within the experimental error (8.7 (6.7) for GG and 21 (19) for GGG). The acceptable domain of parameters for a GGG sequence fully belongs to the corresponding domain for a GG sequence. Thus the domains for GG and GGG base sequences are completely consistent with each other. Therefore we cannot determine parameters λ and b better then using the

"dark" domain for GGG. This information is still sufficient to consider the localization of the hole wavefunction in G_n aggregates.

Since the thermal energy $k_B T \approx 0.026$eV is smaller than at least one of two other characteristic energies of the system (remember that the minimum estimate for the reorganization energy is $\lambda \sim 0.25$eV [19]) we can characterize the wavefunction using the system ground state at coordinates X_i ($i = 1, ... n$) minimizing the ground state energy. In the relevant domain of parameters in Fig. 1 ($\lambda > 0.25$eV) the ground state wavefunction is centered at one of G bases (left or right ones for a GG sequence and the central one for a GGG
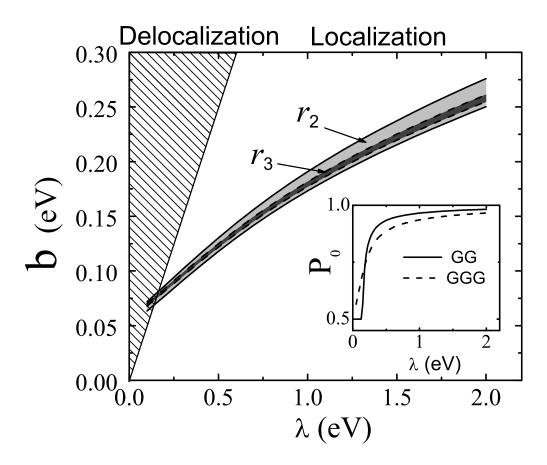

FIG. 1: The domains consistent with the experimental ratios of reaction rates Eq. (3), dark grey for GGG and light grey for GG. Inset shows the fraction of the quantum charge state belonging to the central site (P_0) vs. the reorganization energy λ. Charge transfer integral b is determined in the way to satisfy Eq. (3). We assumed that the system is nearly in its ground state.

sequence). We describe this state by the probability P_0 for the particle in the ground state to be in this central site. This probability can be calculated using the relationship

$$P_0 = \frac{X_i}{\lambda}, \qquad (22)$$

where i is the central site. This equation can be derived as following. Coordinates X_j $(j = 1, ...n)$ in the energy minimum satisfy the condition

$$\frac{\partial \left\langle g \left| \widehat{H} \right| g \right\rangle}{\partial X_j} = 0, \qquad (23)$$

where $|g>$ is the wavefunction of the hole ground state at given coordinates X_j. Using Eq. (5) one can rewrite Eq. (23) as

$$0 = \frac{X_j}{\lambda} - <n_j> . \qquad (24)$$

The average population of the central site i $(<n_i>)$ is equal to the probability P_0 to find the hole there. Accordingly we end up with Eq.(22).

Consider the ground state wavefunction for the hole within the GG sequence. The expression for the hole ground state energy of the GG sequence at arbitrary coordinates X_1, X_2 reads $E_2 = \frac{X_1^2 + X_2^2}{2\lambda} - \frac{X_1 + X_2}{2} - \sqrt{\frac{(X_1 - X_2)^2}{4} + b^2}$. In the regime of interest $2b < \lambda$ (see Fig. 1) the minimum of energy is given by

$$E_{2min} = -\frac{\lambda}{2} - \frac{b^2}{\lambda}, \qquad (25)$$

and it is realized at $X_1 = \lambda/2 \pm \sqrt{(\lambda/2)^2 - b^2} = \lambda - X_2$. Accordingly

$$P_0 = \frac{X_1}{\lambda} = \frac{\lambda + \sqrt{\lambda^2 - 4b^2}}{2\lambda}. \qquad (26)$$

Note that if $2\lambda < b$ the ground state wavefunction is symmetric in the minimum ($X_1 = X_2 = \lambda/2$) and the energy of this state is given by

$$E_{2symm} = -\frac{\lambda}{4} - b. \qquad (27)$$

Consequently $P_0 = 1/2$ (see Fig. 1). In the case of $2b < \lambda$ this symmetric state is the transition state for the charge transfer $G^+G \longleftrightarrow GG^+$ (saddle point in the energy function $E_2(X_1, X_2)$ between the energy minima centered in the first and the second G bases). For GGG complex the probability P_0 that the hole resides in the central site has been evaluated numerically.

Both probabilities P_0 obtained for the λ - b line corresponding to the ratio $r_2 = 7.7$ Eq. (4) are shown in the inset in Fig. 1. It is clear from this graph that for both GG and GGG sequences the wavefunction of the hole is essentially localized within the single G site. For instance at the minimum value of $\lambda \sim 0.25 eV$ we have 85% and 78% of the probability to find the particle in that site for GG and GGG sequences, respectively. These probabilities increase for $\lambda = 1eV$ to 96% and 94%, respectively. Thus we come to the important conclusion that the wavefunction of hole is essentially localized in the single G site for G_n sequences. This conclusion differs from the predictions of previous work [7, 19, 20] where the the quantum state of the hole was represented as a polaron of an intermediate range shared between several base pairs.

In addition to the analysis of the G_2^+ and G_3^+ complexes we have performed the calculations for G_4^+ complexes for experimental verification of our theory. The partition function for G_4^+ complex can be evaluated integrating first the center of mass position similarly to our previous calculations. The predictions for the ratio r_4 are shown in Fig. 2 by the solid line denoted as $n = 4$. The parameters λ and b for calculations were obtained using the experimental condition $r_3(\lambda, b) = 20$ Eq. (3). Using these parameters we calculated the ratio r_2 (solid line marked by $n = 2$) and ratio r_4 (solid line marked by $n = 4$) Eq. (21).

According to Fig. 2 the ratio r_4 does not change much in the whole domain of parameters λ and b satisfying the balance equations Eq. (3). We can estimate it to be around 33 and this is our prediction for the further experimental verification.

B. Understanding of the balance of charge transfer reactions

An impressive consistency between GG and GGG base sequences in Fig. 1 is not accidental coincidence and can be explained by the strong localization of charge wavefunctions. In the regime of strong localization the partition function Z_{n+} for $n \geq 2$ can be represented as the sum of n contributions of energy minima corresponding to wavefunctions centered in all n G sites with coordinates X realizing the corresponding energy minimum $X_i \approx \lambda \gg X_k$, $k \neq i$ for the state centered at site i. Since in the zeroth order approximation in b/λ each quantum state is localized at one site we can neglect the difference in preexponential factors for the case of $b = 0$ and approximate the partial i^{th} contribution to the partition function as $Z_n^i = c^n e^{-\beta E_i^{(0)}}$ (see Eq. (6)), where $E_i^{(0)} = -\lambda/2$ is the energy of the ground state for

coordinates X realizing the local minimum in the absence of electronic coupling ($b = 0$). Second order correction in b to the energy $E_i^{(0)} = -\lambda/2$, which is the first non-vanishing correction, is important because it appears in the exponent $e^{-\beta E}$ and it is multiplied by the large factor β. For two states at the edges this correction adds to the hole energy as $E_1^{(1)} = E_n^{(1)} \approx -\lambda/2 - b^2/\lambda$, which coincides with the ground state energy for GG Eq. (25). For $n - 2$ remaining states the correction to the energy should be doubled because of the addition of contributions of two neighbors so we got $E_i^{(1)} = -\lambda/2 - 2b^2/\lambda$, $1 < i < n$. The
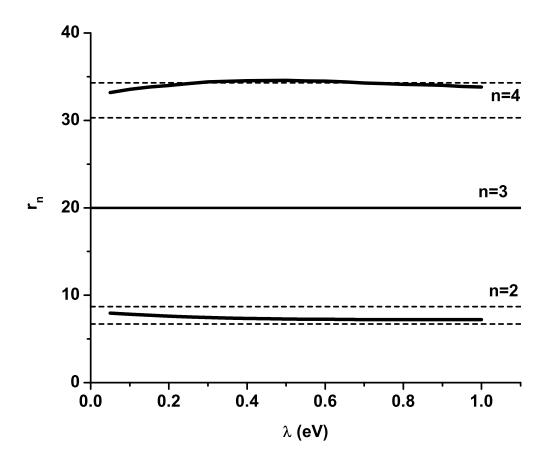

FIG. 2: Predictions for the ratio r_4 using parameters b and λ related by the experimental condition $r_3(b, \lambda) = 20$ Eq. (3). Solid line denoted by $n = 4$ shows the predicted value of r_4 while two dashed lines show upper and lower boundaries for r_4 defined in accordance with Eq. (29). Results for $n = 2$ calculated under the same assumptions are shown for comparison. Upper and lower boundaries for r_2 are given in accordance with the experiment [9], Eq. (3).

15

change in equilibrium environment coordinates can be neglected because near the energy minimum it leads to the effect of order of $(b/\lambda)^4$. Consequently, we can approximate the ratio r_n (cf. Eq. (4)) as

$$r_n = \frac{Z_{n+}Z_1}{Z_nZ_{1+}} \approx 2e^{\beta b^2/\lambda} + (n-2)e^{2\beta b^2/\lambda}. \tag{28}$$

These predictions are compared with the numerical calculations in Fig. 3. One can conclude that perturbation theory works reasonably well in the parameter domain of interest corresponding to $\lambda > 0.25$eV for $n = 2$ and $n = 3$.

In particularly, it follows from Eq. (28) that $(Z_{3+}Z_1/(Z_3Z_{1+}) \approx (Z_{2+}Z_1/(Z_2Z_{1+}) + ((Z_{2+}Z_1/(2Z_2Z_{1+}))^2)$. Indeed, this relationship is satisfied for the experimental values of
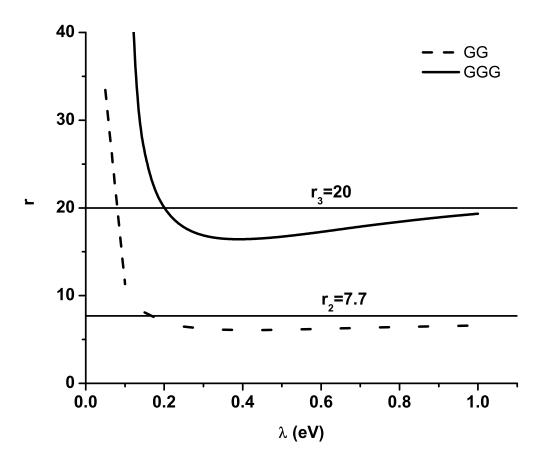

FIG. 3: Comparison of perturbation theory approaches with numerical simulations for ratios r_1 and r_2. Curves show predictions of Eq. (28) for parameters λ and b chosen to satisfy Eqs. (12), (20), respectively ($r_2(\lambda, b) = 7.7$, $r_3(\lambda, b) = 20$).

ratios within the accuracy of the experiment. This explains the consistency of domains for GG and GGG Fig. 1. Using Eq. (28) one can predict that ratios r_n form arithmetic series, which can be expressed as

$$r_n = 7.7 + 12.3 \cdot (n - 2). \tag{29}$$

For balance between G and $GGGG$ sequence we predict the ratio $r_4 = 2r_3 - r_2 = 32.3 \pm 2$. This estimate agrees with our numerical calculations for the $(GGGG)^+$ partition function and reaction rate ratio r_4 (see Fig. 2).

The perturbation theory analysis described above is approximately valid even if the energy of solvent fluctuations Eq. (8) can not be expressed as the bilinear function of coordinates X. One can still define the reorganization energy λ as the energy difference of adjacent G bases in G^+G sequence induced by the environment polarization around the charge localized at the only one (left) base. Then all our conclusions about localization remain valid, but the potential barriers for charge transfer will be larger than in the case of the "harmonic" potential ($\lambda/4$) because higher order nonlinear terms make the barrier sharper.

Since perturbation theory works reasonably well in the parameter domain of interest, we will use it to study more complicated questions of the charge transfer reaction balance in the presence of off-diagonal interaction of solvent polarizations (cf. Eq. (13)) and in the presence of A bases surrounding G_n complexes in experiments [9]. This analysis is reported below in Secs. IV C, IV D, respectively.

C. Perturbation theory in case of off-diagonal correlations of environment polarizations

The most general Hamiltonian of environment in the bilinear approach can be expressed as

$$\widehat{H}_{env} = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} X_i X_j, \tag{30}$$

where the interaction matrix $\widehat{A}$ is defined as the inverse correlation matrix $\widehat{B}$ (cf. Eq. (14))

$$\widehat{A} = \widehat{B}^{-1}, \ B_{ij} = B_{|i-j|} = \frac{< X_i X_j >}{k_B T}. \tag{31}$$

Here average is defined as the thermodynamic average with the Hamiltonian Eq. (30).

Consider the single hole problem using the perturbation theory with respect to the electron transfer integral b. Then in the zeroth order approximation we set $b = 0$ and assume

that the hole is localized in some G base numbered by index k. This problem can be described by the Hamiltonian

$$\widehat{H}_{int} = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} X_i X_j - X_k. \tag{32}$$

The energy minimum is realized at the point where all n derivatives of energy Eq. (32) with respect to X_i $(i = 1, ..n)$ are zeros $\sum_{j=1}^{n} A_{ij} X_j = \delta_{ik}$. The solution to this equation is given by

$$X_i^{(0)} = B_{|i-k|}, \tag{33}$$

and the energy minimum is $E_k^{(0)} = -\frac{B_0}{2}$.

The first non-vanishing correction to the energy due to the small but finite electron transfer integral b appears in the second order of perturbation theory and can be expressed as

$$\delta E_k^{(2)} = -\sum_{i=1}^{n} \frac{b^2 \Delta_{ik}}{X_i^{(0)} - X_k^{(0)}}, \tag{34}$$

where the symbol Δ_{ik} is equal 0 if i and k are non-neighboring G - sites or it is equal 1 for neighboring G sites. Using the zero-point approximation for equilibrium coordinates Eq. (33) we get (cf. Eq. (18))

$$\delta E_k^{(2)} = -z \frac{b^2}{B_0 - B_1} = -z \frac{b^2}{\lambda_*}, \tag{35}$$

where z is the number of neighboring G sites for the center of charge localization. If this site is in the edge of the G_n^+ sequence we get $z = 1$ and if it is in the middle we get $z = 2$. Thus energies of all representative states for G_n^+ sequences can be expressed within the second order of perturbation theory similarly to energies for the diagonal interaction (see Sec. IV B). The preexponential factors in all partition functions in the zeroth order in the electron transfer integral b are given by factors $1/\sqrt{\det A}$, which are all the same in the top and the bottom of ratios r_n. Therefore one can still use Eq. (28) replacing the single site reorganisation energy λ with the charge transfer reorganization energy

$$\lambda_* = B_0 - B_1. \tag{36}$$

This approximation is applicable in the regime of the strong localization. The approximate relationship between parameters of the system can be obtained using Eq. (28) with $n = 3$ where $r_n \approx 20$. Solving this equation for the parameter in exponent we get

$$\frac{b^2}{\lambda_*} \approx 0.033 \text{eV}. \tag{37}$$

18

This is the approximate analytical relationship between the adjacent site reorganization energy and electron transfer integral.

D. Effect of adjacent A bases

In the experiment [9] each G_n complex is always surrounded by A bases. Ionization potential of A base exceeds the one for the G base by approximately 0.5eV [31]. This potential difference is much larger than the estimated value of the electron transfer integral b (see Fig. 1) so we do not expect the crucial effect due to the coupling of G and A bases. The coupling strength of adjacent G and A bases has been estimated in Ref. [18]. For the equilibrium twisting angle $\theta \approx 36^0$ the overlap integral for the GA pair is approximately the same as for the GG pair, while for AG pair it is about two times smaller (data are given for $5' - 3'$ strand). If we take this effect into account using the perturbation theory approach then the ratios r_2 and r_3 take the form

$$r_2 \approx \exp\left(-\frac{\beta b^2}{\lambda_* + \Delta} + \frac{\beta b^2}{\lambda_*}\right) + \exp\left(-\frac{\beta b^2}{4(\lambda_* + \Delta)} + \frac{\beta b^2}{\lambda_*}\right),$$

$$r_3 = r_2 + \exp\left(\frac{2\beta b^2}{\lambda_*} - \frac{\beta b^2}{\lambda_* + \Delta} - \frac{\beta b^2}{4(\lambda_* + \Delta)}\right) \tag{38}$$

These equations can be resolved for λ_* and b. We obtained $\lambda_* = 0.66$eV and $b = 0.18$eV. Clearly this estimate corresponds to the regime of the strong localization. The expectation for the electron transfer integral in the absence of A bases for $\lambda_* = 0.66$eV is $b = 0.15$eV (see Fig. 1) so there is a deviation between two solutions by about 20%. However we cannot take this estimate quite seriously because (1) Our estimate for the electron transfer integral $b = 0.18$eV exceeds the one in Ref. [18] by almost the factor of 2 (2) If we include the experimental error into our estimate then both parameters b and λ_* will not strongly deviate from Eq. (37), while they may vary remarkably compared to the above estimate. Therefore the consideration of A bases seems to be the excess of accuracy.

We believe that the effect of A bases can be clarified experimentally. In the sequence TG_n^+T it will be more justified to ignore surrounding T bases then to ignore A bases in the sequence AG_n^+A studied in Ref. [9]. This is because the ionization potential of T base exceeds that for G base by the factor of 3 more than the one for A base. Therefore the measurements of ratios r_n for G sequences separated by T bases will clarify the effect of neighboring bases and possibly lead to better estimates for b and λ.

One should notice that both the addition of A bases and the correlations of environment coordinates into consideration does not change our prediction that the ratios of reaction rates r_n form the arithmetic series Eq. (29) in the case of the strong localization of charge.

V. CONCLUSION

We have considered the quantum state of the positive charge (hole) in poly-G - poly-C base sequence. We have studied a very simple but quite general model of G_n sequence characterized by two parameters including electron transfer integral b and reorganization energy λ. Our choice of parameters b and λ was determined by the the comparison of theory and experiment for the ratios r_n ($n = 2, 3$) of the charge transfer rate between single G and G_n complexes. It turns out that the agreement with the experimental data for the ratios r_n Eq. (4), Ref. [9] can be attained only assuming the strong localization of charge within a single G-base. The charge in DNA then behaves as a small polaron with the size less than the interbase distance. Charge hopping takes place due to rare environment fluctuations. Charge transfer rate is determined by the probability of such fluctuation and should obey the Arrhenius law. Based on our theory we predict that ratios of charge transfer rates form the arithmetic series $r_n = 7.7 + 12.3 \cdot (n - 2)$. We propose the experiment measuring the ratios of rates for $n \geq 4$ as a direct verification of our theory. However we cannot uniquely identify specific values of the electron transfer integral b and the reorganization energy λ. Measurements of similar ratios in TG_nT sequences can be used to verify the theory and to make better parameter estimates.

In addition one can attempt to determine these parameters by measuring the temperature dependence of the charge transfer rate through poly-G - poly-C base sequence. We expect that this temperature dependence can be described by the Arrhenius law with the activation energy defined by the difference of charge symmetric transition state energy within $(GG)^+$ base sequence Eq. (27) and the charge ground state energy for $(GG)^+$ state Eq. (25) $E_A = \lambda/4 - b + b^2/\lambda$ [32]. Since another relationship between b and λ is known from the ratio of charge transfer rates r_2 and r_3 (see Fig. 1) this information will be sufficient to find both parameters. Note that the temperature range where such dependence can be measured is quite narrow (380K$< T < $410K) so it will be hard to prove the Arrhenius law based on such experimental data. Therefore it can be only assumption that this law is

indeed applicable. Despite of the narrow temperature range the remarkable change in the charge transfer rate is expected with growing the temperature from its minimum 380K to its maximum 410K. For instance if $\lambda = 0.3$eV this rate will increase by the factor 1.5 with increasing temperature, while if $\lambda = 1$eV then the rate will increase by the factor 3.8.

---

[1] C. J. Murphy, M. R. Arkin, Y. Jenkins, N. D. Ghatlia, S. H. Bossmann, N. J. Turro, and J. K. Barton, *Science,* Vol. **262**, 1025 (1993).

[2] M. R. Arkin, E. D.A. Stemp, R. E. Holmlin, J. K. Barton, A. Hoermann, E. J.C. Olson, P. F. Barbara, *Science*, **273,** 475 (1996).

[3] D. B. Hall, R. E. Holmlin and J. K. Barton, *Nature*, **382**, 731 (1996).

[4] E. Meggers, M. E. Michel-Beyerle, B. Giese, J. Amer. Chem. Soc. **120**, 12950 (1998).

[5] P. T. Henderson, D. Jones, G. Hampikian G, Y. Z. Kan, G. B. Schuster, Proc. Natl. Acad. Sci. **96**, 8353 (1999).

[6] S. R. Rajski, B. A. Jackson, and J. K. Barton, Mutat. Res. **447**, 49 (2000).

[7] C. T. Shih, S. Roche, R. A. Romer, Phys. Rev. Lett. **1**, 018105 (2008).

[8] D. Porath, A. Bezryadin, S. de Vries and C. Dekker, *Nature*, **403**, 635 (2000); T. Takada, K.Kawai, M. Fujitsuka, T. Majima, Proc. Natl. Acad. Sci. **101**, 14002 (2004).

[9] F. D. Lewis, X. Liu, J. Liu, S. E. Miller, R. T. Hayes and M. R. Wasilewski, *Nature*, **406**, 51-53 (2000); F. D. Lewis, X. Liu, J. Liu, R. T. Hayes and M. R. Wasilewski, *J.Am. Chem. Soc.*, **122**, 12037 (2000).

[10] H. Sugiyama, I. Saito*J. Am. Chem. Soc.*, **118**, 7063 (1996).

[11] All quantum states in one-dimensional random system are localized at the scale $l \sim a(b/\Delta)^2$ where a is the interatomic distance, b is overlap integral and Δ is disordering in energy, see e. g. B. Shapiro, Phys. Rev. Lett. **48**, 823 (1982).

[12] M. Unge, S. Stafstrom, Nano Lett. **3**, 1417 (2003).

[13] M. Bixon, B. Giese, S. Wessely, T. Langenbacher, M. E. Michel-Beyerle, J. Jortner, Proc. Natl. Acad. Sci. **96**, 11713 (1999).

[14] Y. A. Berlin, A. L. Burin, M. A. Ratner, J. Phys. Chem. A **104**, 443 (2000).

[15] I. V. Kurnikov, G.S.M. Tong, M. Madrid, D. Beratan, *J. Phys. Chem. B*, **106**, 7 (2002).

[16] Y. A. Berlin, I. V. Kurnikov, D. Beratan, M. A. Ratner, A. L. Burin, Topics in Current Chemistry **237**, 1 (2004).

[17] D. LeBard, M. Lilichenko, D. Matyushov, Y. A. Berlin and M. Ratner, *J. Phys. Chem. B*, **107**, 14509 (2003).

[18] A. A. Voityuk, N. Rösch, M. Bixon, J. Jortner, J. Phys. Chem. **104**, 9740 (2000); Kh. Siriwong, A. A. Voityuk, M. D. Newton, N. Rösch, *J. Phys. Chem. B*, **107**, 14509 (2003); A. A. Voityuk, Chem. Phys. Lett. **451**, 153 (2008).

[19] J. A. Berashevich, T. Chakraborty, J. Phys. Chem. B **111**, 13465 (2007).

[20] E. M. Conwell, D. M. Basko *J. Am. Chem. Soc.* **123**, 11441 (2001); E. M. Conwell, S. M. Bloch, P. M. McLaughlin, D. M. Basko, *J. Am. Chem. Soc.* **129**, 9175 (2007); X. F. Wang and T. Chakraborty, *Phys. Rev. Lett.,* **97**, 106602 (2006)

[21] R. A. Marcus, *Rev. Mod. Phys.*, **65**, 599 (1993).

[22] J. N. Onuchic, D. N. Beratan, J. J. Hopfield, J. Phys. Chem. **90**, 3707 (1986); A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fischer, W. Zwerger, Rev. Mod. Phys. **59**, 1 (1987).

[23] A. L. Burin, Yu. A. Berlin, M. A. Ratner, J. Phys. Chem. A **105**, 2652 (2001)).

[24] A. L. Burin, I. Y. Polishchuk, P. Fulde, Y. Sereda, Phys. Rev. Lett. **96**, 025505 (2006).

[25] A. A. Voityuk, K. Siriwong, N. Rosch, Angew. Chem. Int. **43**, 624 (2004).

[26] M. Buttiker, R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982); Y. Imry, R. Landauer, Rev. Mod. Phys. **71**, S306 (1999).

[27] A. Nitzan, J. Jortner, J. Wilkie, A. L. Burin, M. A. Ratner, J. Phys. Chem. B 104, 5661 (2000).

[28] A. A. Voityuk, J. Chem. Phys. **128**, 045104 (2008).

[29] B. Giese, J. Amaudrut, A. K. Kohler, M. Spormann, S. Wessely, Nature **412**, 318 (2001).

[30] Yu. A. Berlin, A. L. Burin, M. A. Ratner, Chemical Physics 275, 61 (2002).

[31] A. A. Voityuk, J. Jortner, M. Bixon, N. Rösch, Chemical Physics Letters **324**.430 (2000).

[32] We expect that the charge transfer between adjacent guanines occurs adiabatically. The adiabatic regime takes place, when the Landau-Zener criterion $b^2 > \hbar < |dE/dt| >$ is satisfied, where b is the electron transfer integral and dE/dt is the characteristic rate of energy change [21]. For crude estimate of dE/dt one can assume that the energy changes are due to the

solvent (water) and use $dE/dt \sim \lambda/\tau$, where $\tau \sim 8\text{ps}$ is the characteristic relaxation rate of the water polarization (e. g. J. B. Hasted, Liquid water: Dielectric properties, in Water A comprehensive treatise, Vol 1, Ed. F. Franks (Plenum Press, New York, 1972) pp. 255-309) and $\lambda \sim 0.5\text{eV}$ is the reorganization energy for the charge transfer. Then for the electron transfer integral $b > b_c \approx 0.006\text{eV}$ the charge transfer takes place adiabatically. To our knowledge majority of theoretical estimates of b exceed this threshold value b_c by at least an order of magnitude (e. g. [10, 18]), which justifies our expectation.